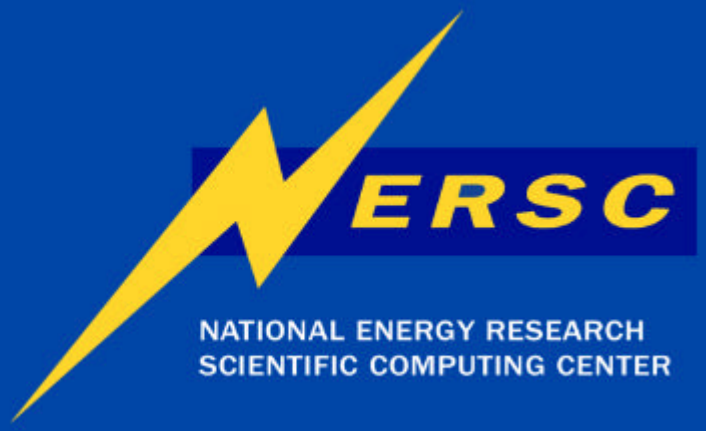


NERSC Resources

Francesca Verdier
Associate Manager, NERSC Services
fverdier@lbl.gov

Workshop on high end computing for nuclear fission science and engineering
February 23, 2006

NERSC Mission

The mission of the National Energy Research Scientific Computing Center (NERSC) is to accelerate the pace of scientific discovery by providing high performance computing, information, data, and communications services for research sponsored by the DOE Office of Science (SC).

Office of Science
U.S. DEPARTMENT OF ENERGY

Science-Driven Computing

NERSC's role is to enable new science

Office of Science
U.S. DEPARTMENT OF ENERGY

NERSC Center Overview

- Funded by DOE, annual budget $38M, about 60 staff
- Enhances open, unclassified, basic research
- Supports about 2,500 users working on about 300 projects
- Located next to University of California, Berkeley campus (UCB)
- Close collaborations with UCB and the Computational Research Division at LBNL in computer science and computational science

NERSC Strengths

- **Large scale science**
  - Both computational and data intensive science
  - Focus on applications that are able to use a large fraction of available resources (CPU, memory, disk)
- **Effective systems management and support**
  - High MTBF; high availability (98-99.9%)
  - Robust scientific environment (tools, libraries, applications)
  - Full Services (scientific consulting, application tuning, network analysis)
  - Responsive to user needs (specialized consulting, software installations, queue priorities, etc)
- **Center Balance**
  - Storage, networking, and visualization to match computational resources
- **Computer Security**
  - Balance security with user effectiveness

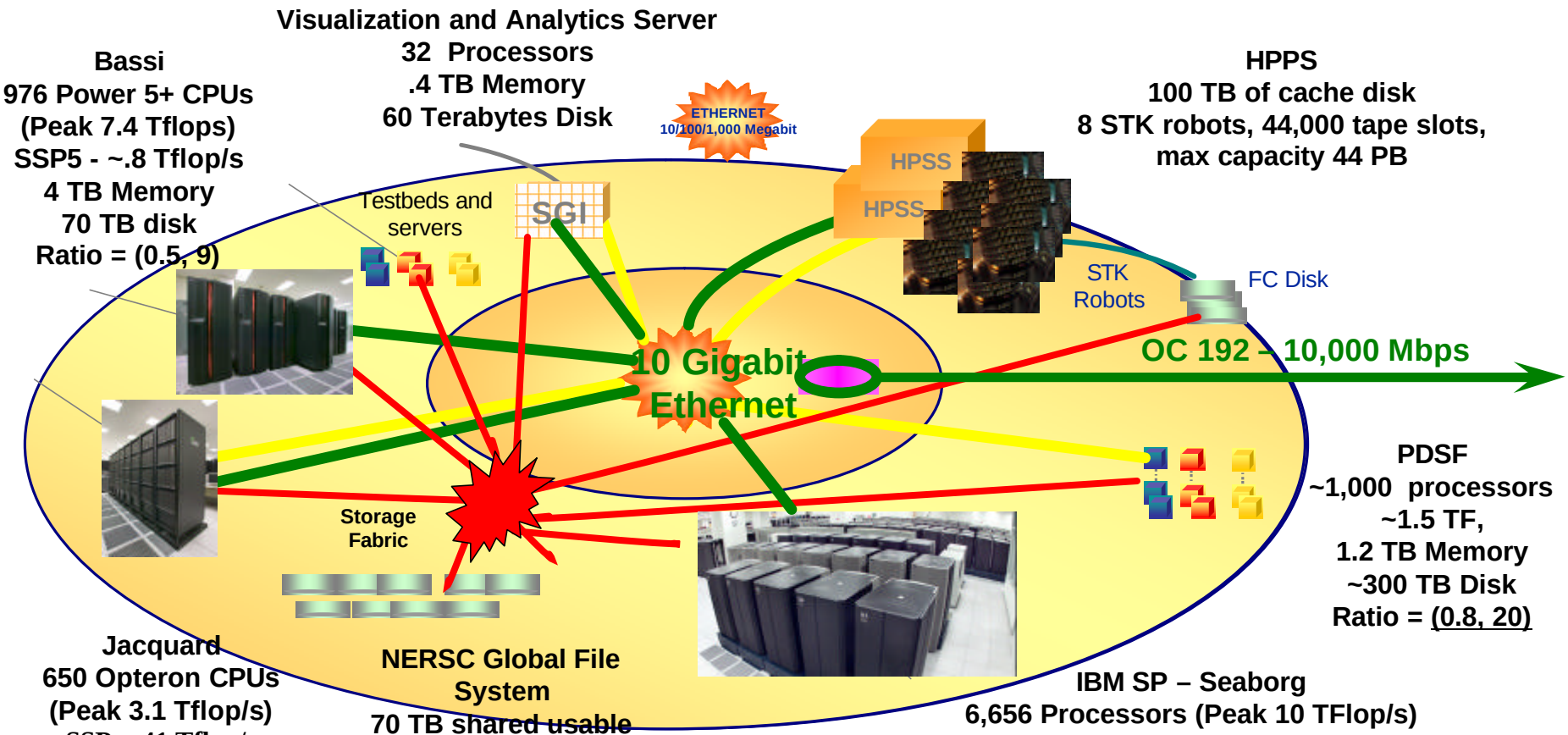
Science-Driven Computing Strategy 2006 -2010

Office of Science
U.S. DEPARTMENT OF ENERGY

NERSC Science-Driven Systems

- **Five major areas**
  - **Computational systems**
  - **NERSC Global Filesystem**
  - **Mass storage systems**
  - **Network**
  - **Analytics and visualization systems**
- **The balance of the entire Center is determined by**
  - **Requirements that evolve from increased computational capability**
  - **Independent requirements for other resources, such as handling experimental data, analytics requirements**

NERSC Configuration
January 2006



Ratio = (RAM Bytes per Flop, Disk Bytes per Flop)

Office of Science
U.S. DEPARTMENT OF ENERGY

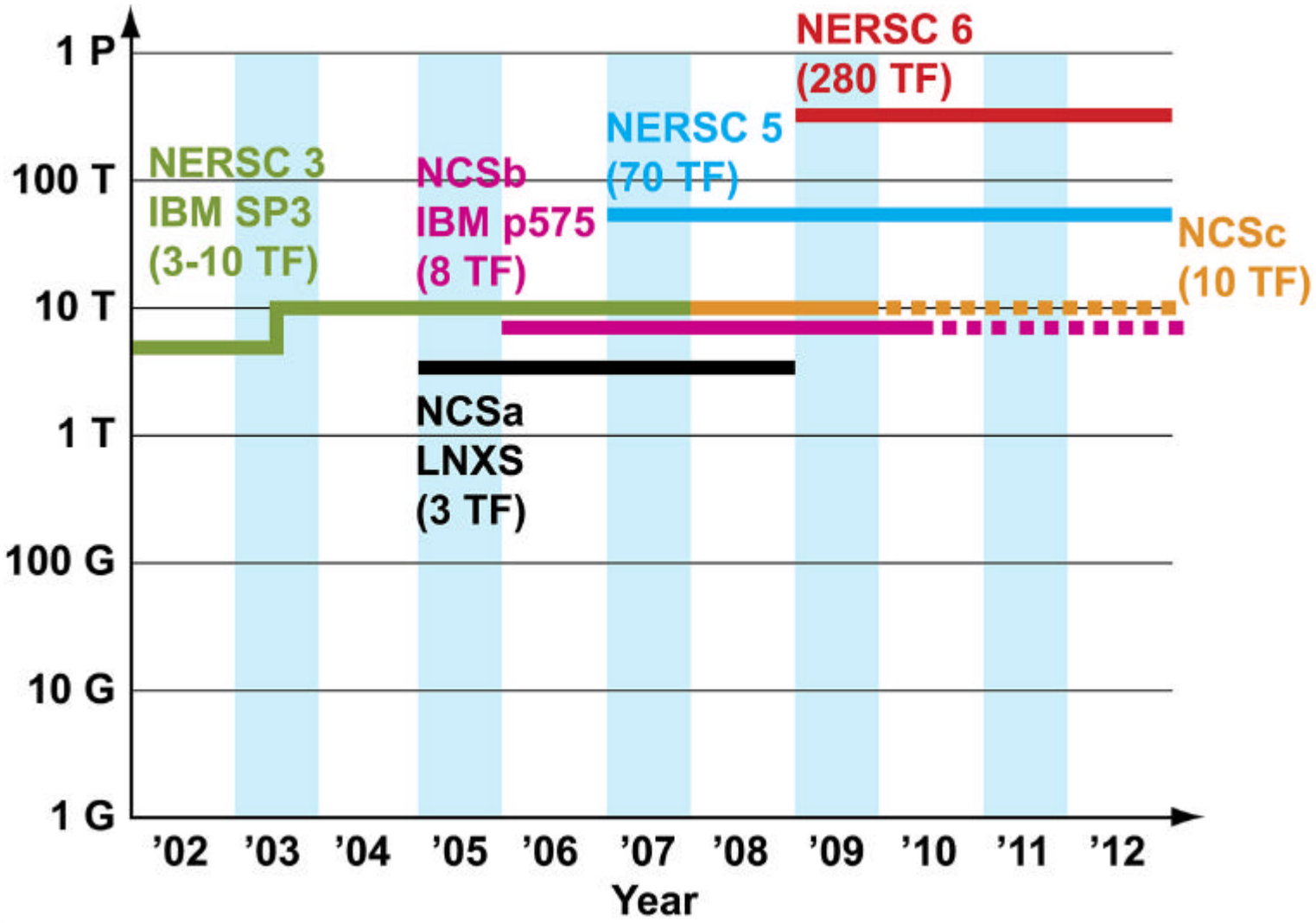
Bassi POWER5 System

- 111 IBM p575 POWER5 compute nodes:
  - 8 processors per node
  - 32 GB shared memory per node
  - 1.9 GHz POWER 5 processors (7.6 Gflop/s peak)
- 3.55 TB aggregate memory (on compute nodes)
- 70 TB shared GPFS disk
- High bandwidth, low latency switch (IBM Federation)
  - ~4.7 microseconds MPI point-to-point latency
  - ~1.5 GB/sec unidirectional MPI bandwidth

Jacquard Linux Cluster

- 320 Opteron compute nodes:
  - 2 processors per node
  - 6 GB shared memory per node
  - 2.2 GHz Opteron processors (4.4 Gfop/s peak)
- 1.92 TB aggregate memory (on compute nodes)
- 30 TB shared GPFS disk
- High bandwidth, low latency InfiniBand interconnect
  - ~4.5 microseconds MPI point-to-point latency
  - ~620 MB/sec unidirectional MPI bandwidth

NERSC Systems Roadmap



Office of Science
U.S. DEPARTMENT OF ENERGY

Science-Driven Services

- Provide the entire range of services from high-quality operations to direct scientific support
- Enable a broad range of scientists to effectively use NERSC in their research
- Concentrate on
  - Scaling to large numbers of processors
  - Data management and analysis
  - Supporting multidisciplinary computational science teams

Science Driven Service Strategies

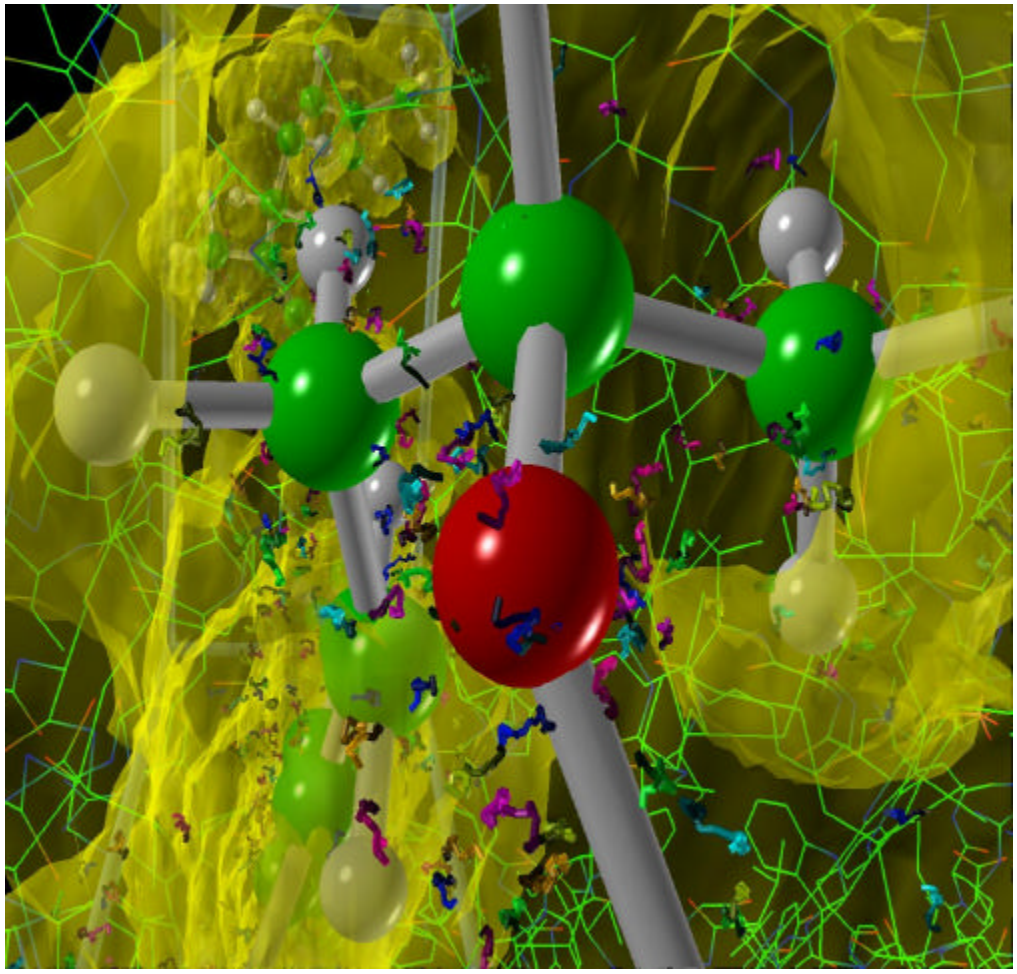
- Make NERSC the best center for computational scientists
- Address the unique issues of using large-scale systems
- Drive Center changes based on user interactions
- Ensure that computing systems and software are reliable (secure, stable, and well performing)

Make NERSC the best center for computational scientists

Photosynthesis INCITE Project

PI: William Lester, UC Berkeley

- MPI tuning: 15-40% less MPI time
- Quantum Monte Carlo load balancing: 256 to 4,096 procs
- More efficient algorithm for random walk procedure
- Wrote parallel HDF5 I/O layer
- Used AVS/Express to visualize molecules and electron trajectories
- "Visualization has provided us with modes of presenting our work beyond our wildest imagination"
- "We have benefited enormously from the support of NERSC staff"

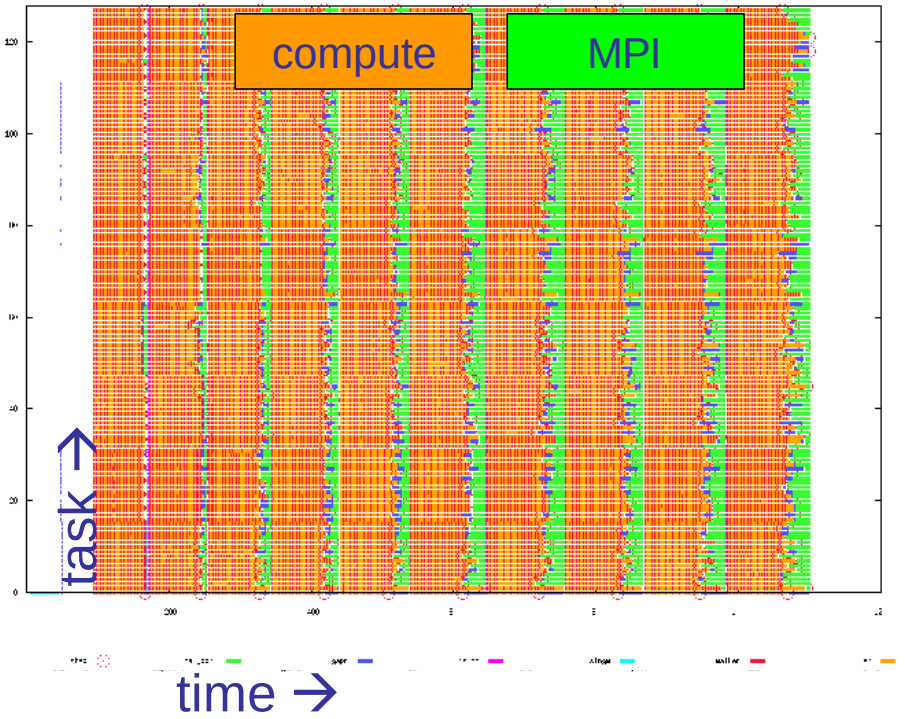


Office of Science
U.S. DEPARTMENT OF ENERGY

Address the unique issues of using large-scale systems

Code Performance Tuning

- Analysis of Parallel Performance
  - Detailed view of how the code performs
- Code Performance Tuning
  - MPI Tuning
  - Scaling and Job structuring
  - Use optimally tuned math libraries
- Parallel I/O strategies
  - Tuning for seaborg
  - Use optimal I/O libraries



Drive Center changes based on user interactions

- NERSC Users' Group
- User Requirements
  - Procurement and Visualization Greenbooks
- User Surveys
  - Annual User Survey
  - Annual Allocations Request Form
  - Early user feedback questionnaires
- User Feedback
  - Consulting interactions: users provide feedback on issues that affect them
  - Scientific meetings and workshops: Interact with NERSC user attendees
  - Informal interactions (email exchanges, personal exchanges)

Ensure that computing systems and software are reliable

Keep Systems Secure but Usable

- Implement effective computer security measures while maintaining the ability of NERSC users to do science
- Maintain open, unencumbered access at high performance rates
- Spring 2004 security incident:
  - Over 2000 university and government systems were root compromised
  - Many centers shut down for weeks
  - Due to our security practices, NERSC remained up and secure

Science-Driven Analytics

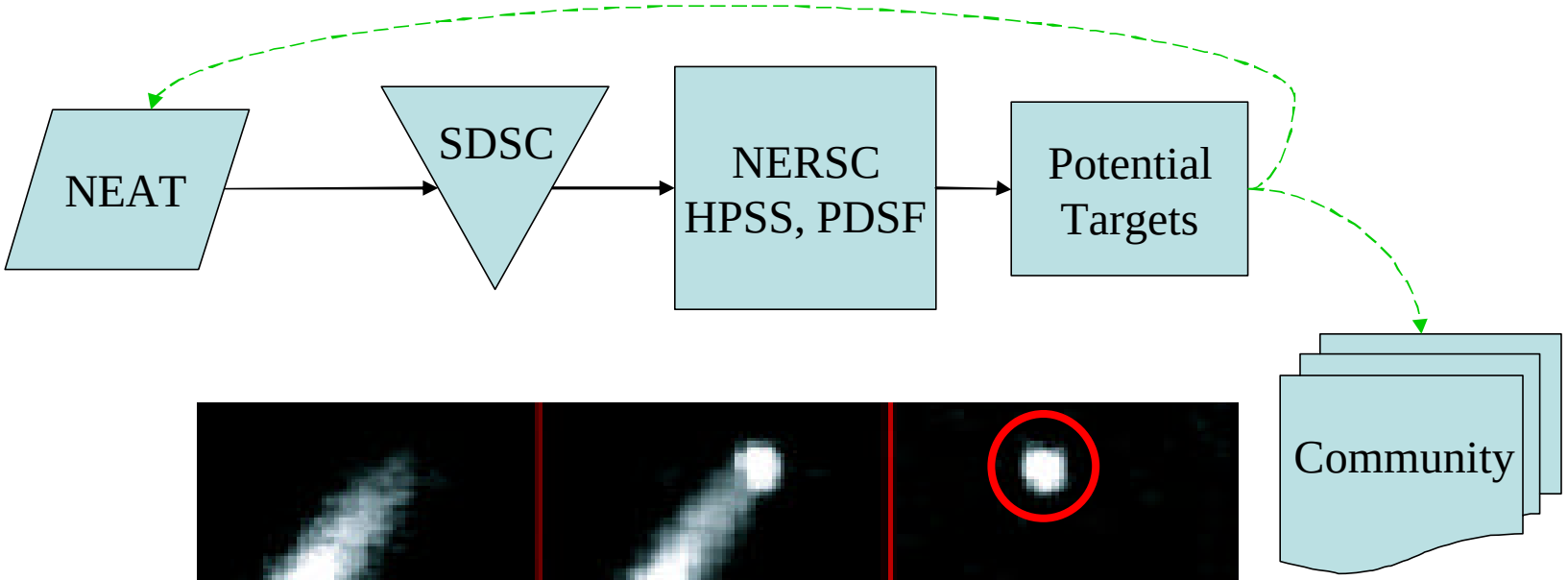
- Provide architectural and systems enhancements and services to more closely integrate computational and storage resources
- Provide scientists with new tools to effectively manipulate, visualize and analyze the huge data sets from both simulations and experiments

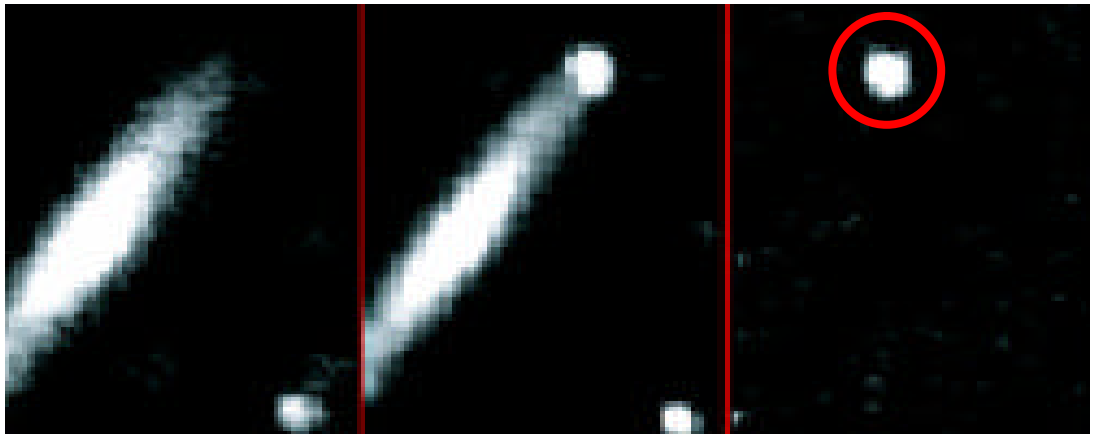
What is Analytics?

- **Science of reasoning**
  - **Generate insight and understanding from large, complex, disparate, sometimes conflicting data**
- **NERSC Analytics Program Elements**
  - **Users and their data analysis problems!**
  - **Visualization**
  - **Analysis**
  - **Data management**
  - **Workflow management**

Analytics Example

- **Nearby Supernova Factory data analysis pipeline:**

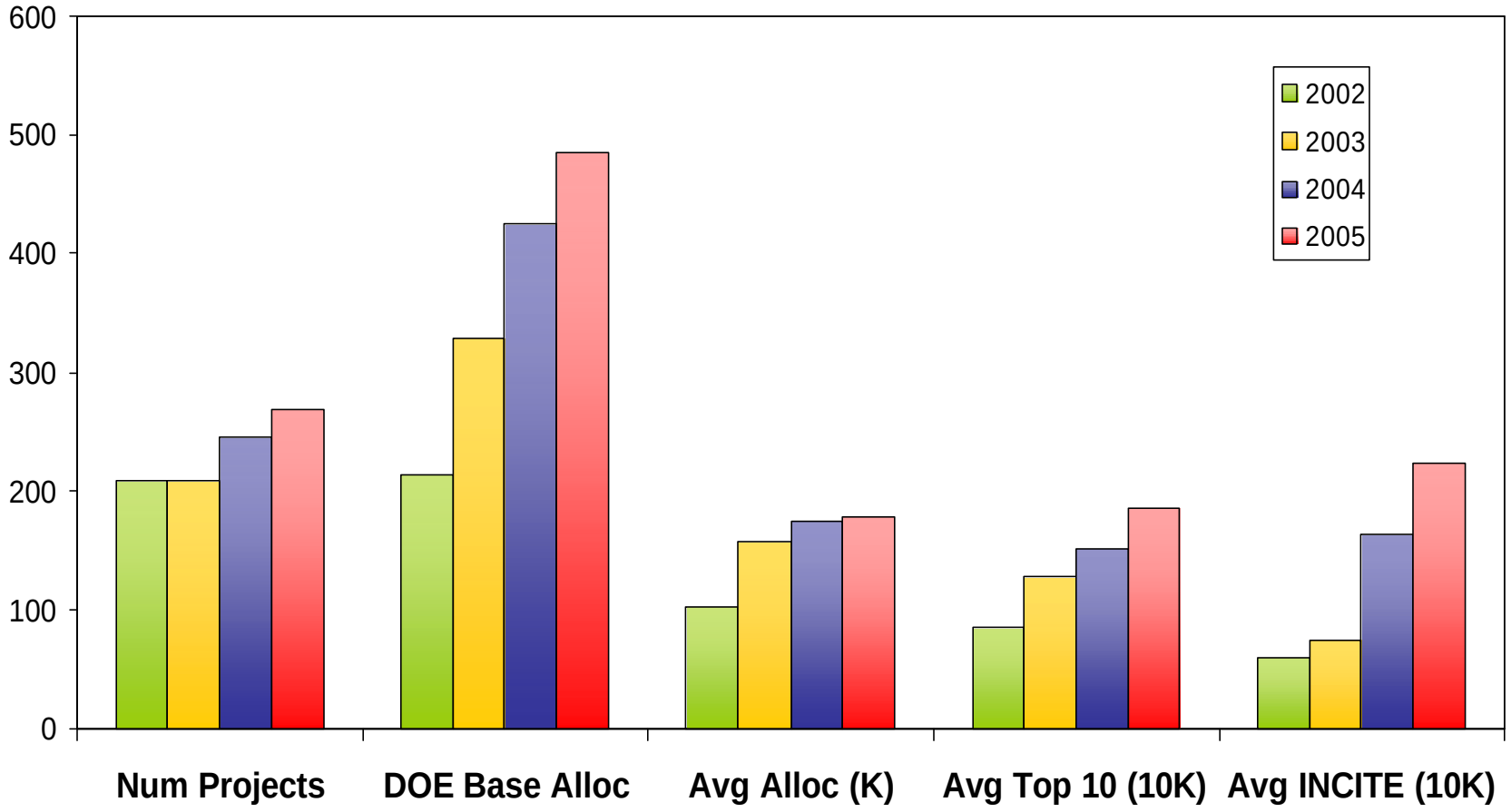




NERSC User Demographics

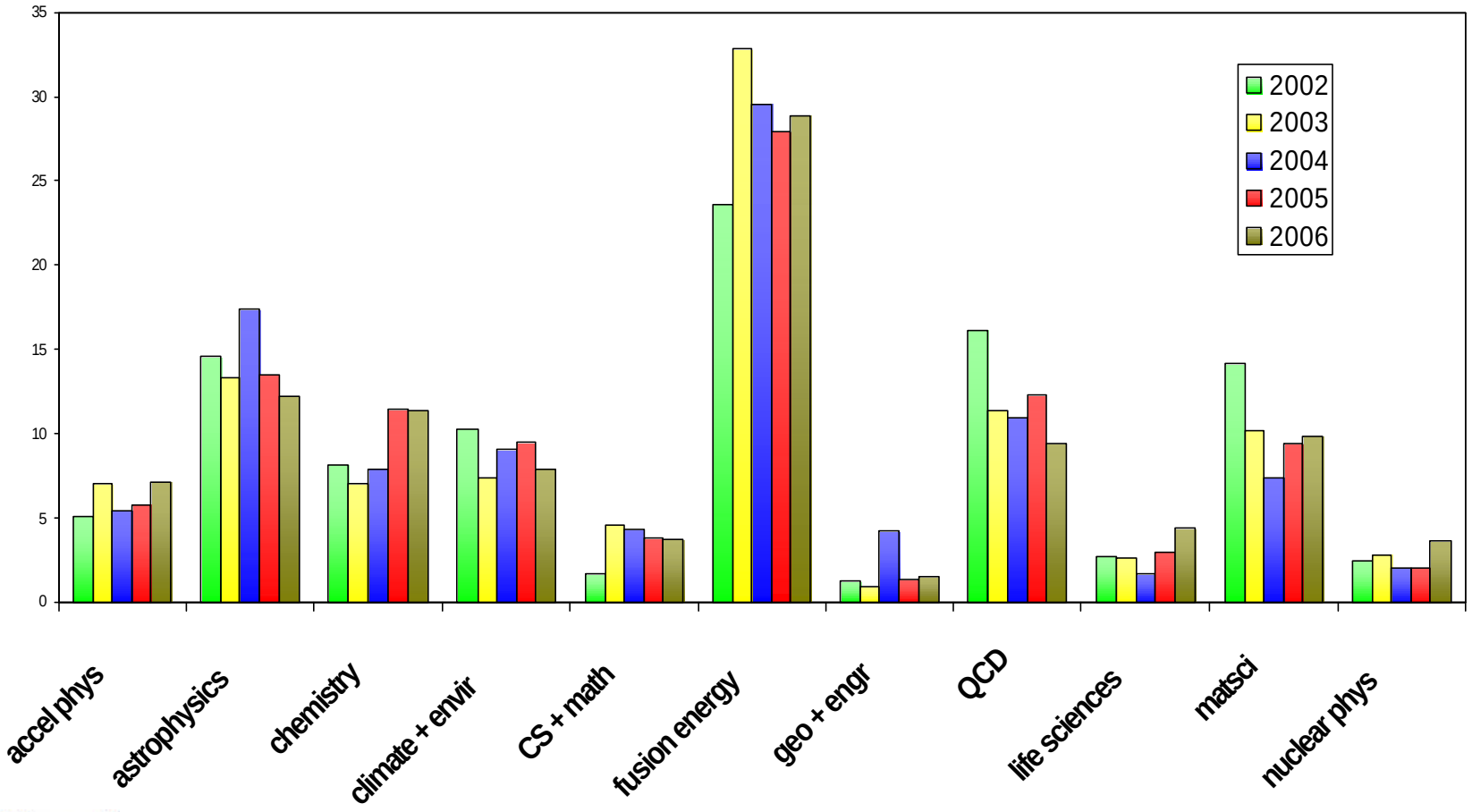
- **Allocations and Projects**
  - Average allocations for the largest projects are increasing the fastest
  - Growing number of projects
    - Average allocation size is stagnant; hard for projects to scale
  - Fusion dominates the computational allocations
  - Physical sciences dominate the storage allocations
  - Fission is not yet represented at NERSC
- **Users:**
  - Growing number of users, especially Storage users
  - Need for more data analysis (analytics)
  - Need for more data related infrastructure

MPP Allocation Profile: INCITE Allocations Growing the Fastest



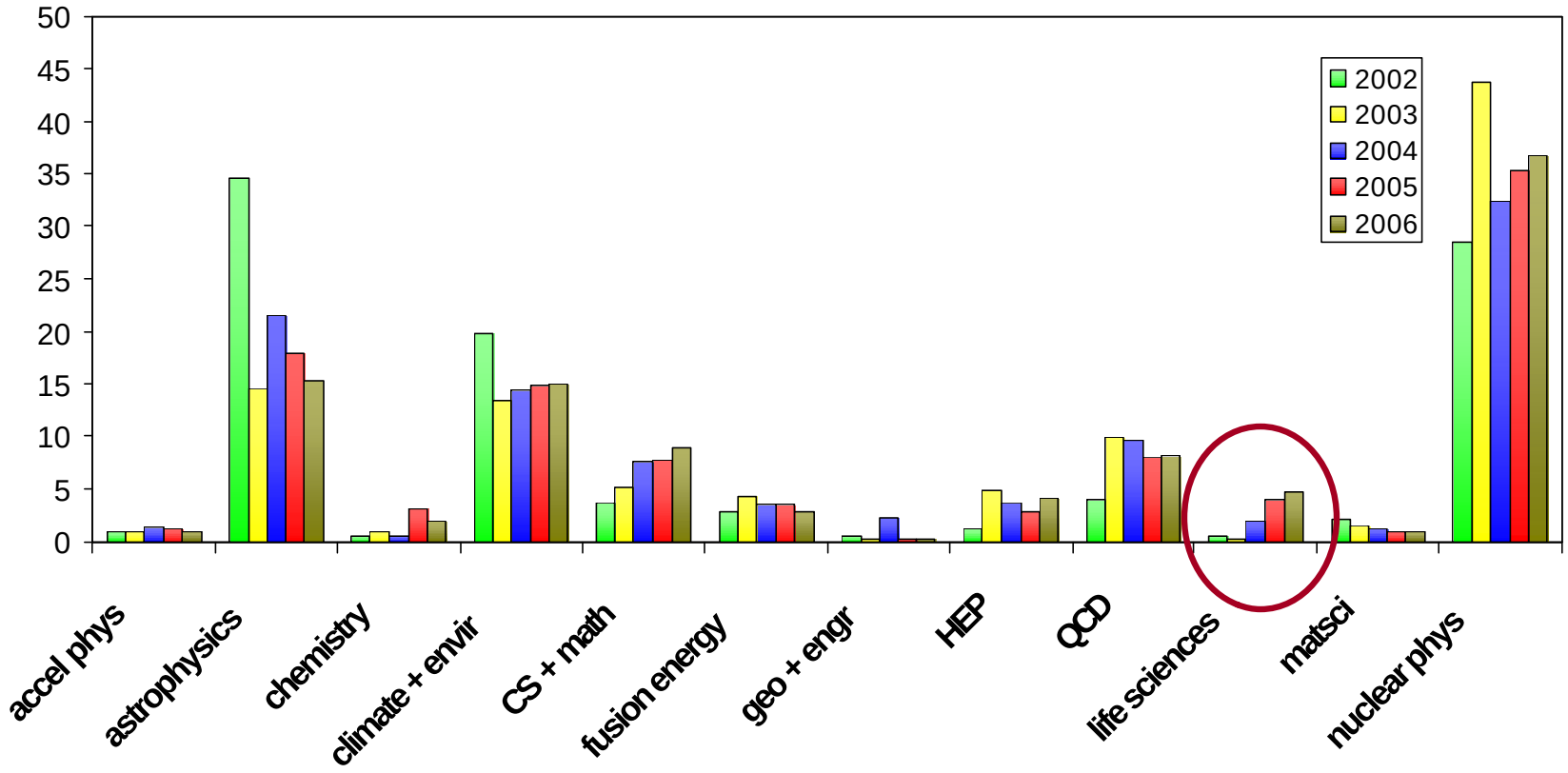
Fusion Dominates MPP Use
(Nuclear fission not currently represented)

Percent of MPP Allocation to Science Areas



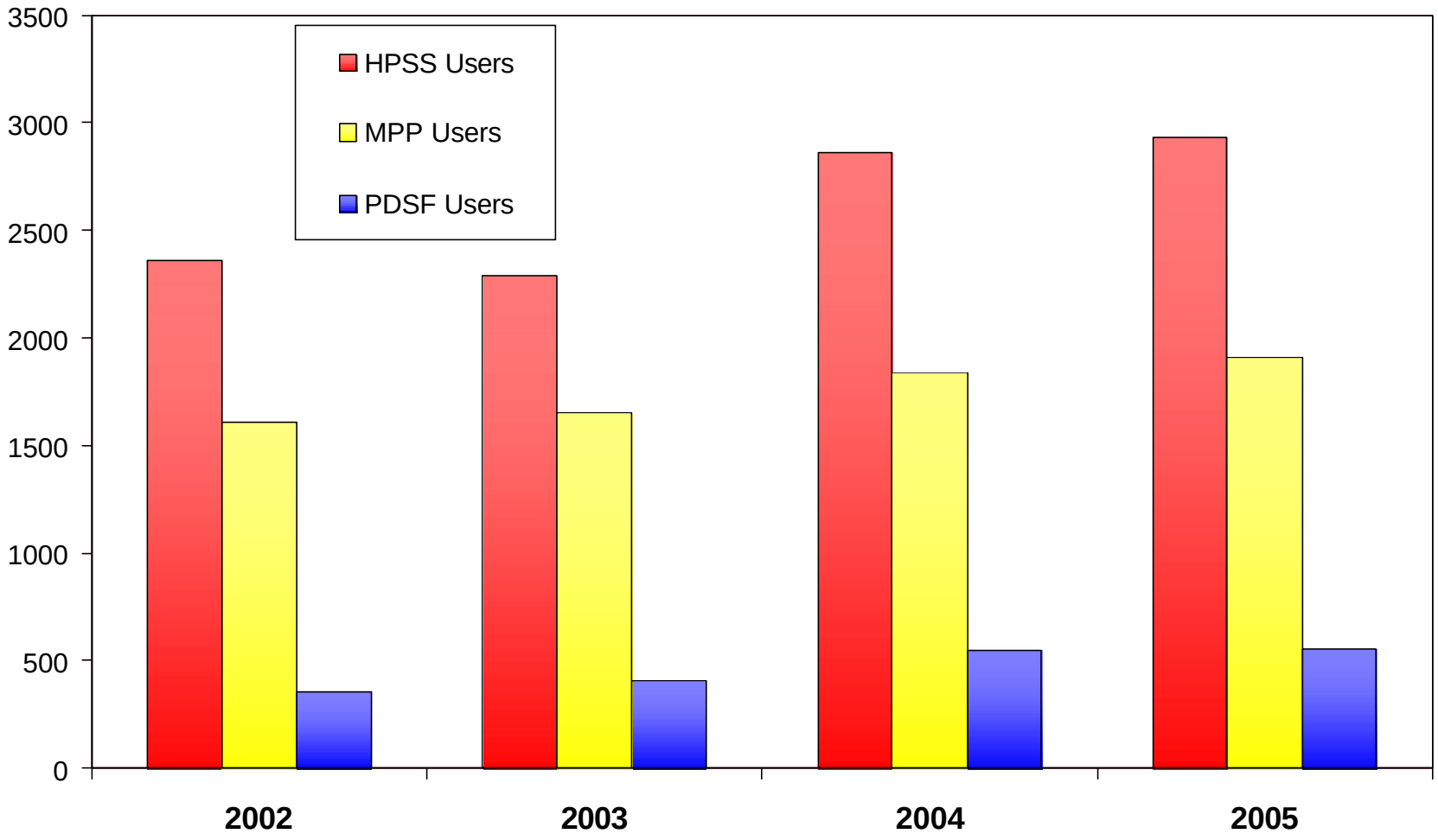
Physical Sciences Dominate Storage Use; Life Sciences Taking Off

Percent of HPSS Allocation to Science Areas



Office of Science
U.S. DEPARTMENT OF ENERGY

Increasing Numbers of Storage Users = Greater Focus on Data Management



NERSC and Nuclear Fission

- Need to understand your requirements
- Computational Fluid Dynamics and Transport
  - Could install new software at NERSC to meet fission needs
  - Very active nuclear fusion transport & MHD user communities
- Computational Materials Science
  - Very active user community: 52 projects
  - Large software base: VASP, FLAPW, Paratec, PEscan, PEtot, PWscf, Siesta, WIEN2K, ABINIT, AL_CMD, CHARMM, DL_POLY, Gaussian, LAMMPS, LSMS, Moldy, NAMD, NWChem, TBMD
- Computational Actinide Chemistry
  - Active DOE user community
  - NWchem, GAMESS (special LAPI version for the POWER architecture)
  - Needs low switch latency (Bassi and Jacquard)
- Nuclear engineering; reactor and cross-section physics
  - Not currently represented at NERSC
  - Overlap with accelerator design?

FY07 Office of Science Allocations

- **FY07 Allocations**
  - Early April: INCITE opens http://hpc.science.doe.gov/
  - May 1: ORNL Leadership / NERSC
  - June 2: INCITE closes
  - June 15: ORNL Leadership closes
  - Mid July: NERSC closes
  - Mid Sept: all awards announced
- **NERSC Startup allocations**
  - Can apply at any time
  - http://www.nersc.gov/nusers/accounts/allocations/
- **PNNL EMSL: http://www.emsl.pnl.gov/using-emsl/process.shtml**
- **ORNL pilot projects: contact Doug Kothe, kothe@ornl.gov**